DOI:10.1111/hpb.12294 HPB

ORIGINAL ARTICLE

Pancreatic fistulae after a pancreatico-duodenectomy: are pancreatico-gastrostomies safer than pancreatico-jejunostomies? An expertise-based trial and propensity-score adjusted analysis

Jad Abou Khalil[1], Nancy Mayo[2], Sinziana Dumitra[1], Mohammed Jamal[3], Prosanto Chaudhury[1], Peter Metrakos[1] & Jeffrey Barkun[1]

[1]Department of General Surgery, [2]Division of Clinical Epidemiology, McGill University, Montreal, QC, Canada, and [3]Department of Surgery, Kuwait University, Kuwait City, Kuwait

Abstract

Background: A pancreatic fistula (PF) is a major contributor to morbidity and mortality after a pancreaticoduodenectomy (PD). There remains debate as to whether re-establishing pancreaticoenteric continuity by a pancreatico-gastrostomy (PG) can decrease the risk of a PF and complications compared with a pancreatico-jejunostomy (PJ). The outcomes of patients undergoing these reconstructions after a PD were compared.

Method: Patients undergoing a PD between 1999 and 2011 were selected from a prospective database and having undergone either a PG or PJ reconstruction. A propensity-score adjusted multivariate logistic regression was performed to identify the effect of surgical technique on outcomes of PF, delayed gastric emptying (DGE) and total complications.

Results: Twenty-three out of 103 and 20 out of 103 ($P = 0.49$) patients had PF and 74 out of 103 and 55 out of 103 patients had all-grades DGE in the PG and PJ groups, respectively ($P = 0.02$). The groups did not differ with regards to Clavien–Dindo grade of complications ($P = 0.29$) but did differ with regards to the Comprehensive Complication Index (CCI) (38.4 versus 31.4 for PG versus PG, respectively, $P = 0.02$.) Propensity-score adjusted multivariate analysis showed no effect of PG on PF ($P = 0.89$), DGE grades B/C ($P = 0.9$) or CCI ($P = 0.41$). There remained an effect on all-grades of DGE ($P = 0.012$.)

Discussion: Patients undergoing PG reconstruction had a similar rate of PF as those undergoing a PJ after a PD.

Received 4 March 2014; accepted 10 May 2014

Correspondence

Jad Abou Khalil, 687 Pine Avenue West, Montreal, QC H3A1A1, Canada. Tel: +1 (514) 294-6867. Fax: +1 (514) 288-8196. E-mail: jad.aboukhalil@mail.mcgill.ca

Introduction

Pancreatic cancer is now the 10th most common cancer and the 4th cause of cancer death in both men and women in North America.[1] Surgical resection provides the only opportunity for a cure, but only 10% of patients with pancreatic adenocarcinoma are surgical candidates at the time of diagnosis.[2] In addition to the complications of abdominal surgery, pancreatic fistulae (PF) weigh heavily on the morbidity and mortality burden incurred by patients after a pancreaticoduodenectomy (PD).[3] The traditional approach uses pancreatico-jejunostomy (PJ) and many surgeons have proposed variations thereof, such as dunking, Peng's binding and the Blumgart approach.[4,5] However, pancreatico-gastrostomies (PG) have gained favour with some surgeons, as meta-analysis of retrospective series have suggested superiority with regards to PF.[6] However these studies are small, heterogeneous and did not adjust for known risk factors of PF.[7] Most randomized trials to date have failed to convincingly support the meta-analysis findings,[8] thus there remains debate regarding the best approach. Ongoing randomized trials are attempting to answer the question.[9] Although conventional randomized designs are generally less

This study was presented at the Annual Meeting of the AHPBA, 19-23 February 2014, Miami, Florida.

prone to bias, the internal and external validity of randomized trials examining this particular issue may be compromised. The main issue is the randomization process which forces a surgeon to perform both repairs such that there may not be true equipoise in that there is a belief by the surgeon that one option is superior. An informal survey at the 2014 American Association for Hepato-Pancreao-Biliary Surgeons meeting revealed that, respectively, 4% and 2% of surgeons stated they 'always' or 'frequently' use PG, with 62% stating they never use PG and 23% only using it occasionally (conference proceedings, unpublished data). Deveraux and Bandahri have suggested an expertise-based design where patients are randomized to a surgeon performing their 'best' operation.[10]

At our institution, referred patients are assigned non-purposefully to surgeons each having a strong personal preference for either PG or PJ, and each operating well beyond the procedural learning curve for their 'favoured' repair – this unique setting is not unlike that expertise-based trial, wherein each procedure is performed expertly and proficiently, with other aspects of patient care (pre-operative workup and post-operative management) being identical. Our goal was to examine the morbidity experience of PG or PJ after a PD at our institution and to use a propensity-score to adjust for differences between the groups and mitigate confounding.

Patients and methods

Setting, patient selection and group allocation

Patients having undergone a PD from 1999 to 2011 for any indication at the McGill University Health Centers (MUHC) were identified from a prospectively collected database and operating room records. Patients operated by surgeons with a surgical volume of less than five were excluded. Medical records were reviewed to supplement the database after obtaining approval from the hospital ethics review board. The MUHC is a tertiary care centre where three experienced surgeons performing a PD have a high personal preference for PJ or PG reconstruction, and where patients are assigned to surgeons based on each surgeon's waitlist availabilities. The surgeons are fellowship-trained hepatobiliary surgeons with an equivalent experience and more than 150 pancreatectomies documented before 1999. Surgeons did not alter their preferred technique based on pancreatic texture or duct diameter. The hospital, in the city of Montreal, is one of three provincial referral centres for hepato-pancreatico-biliary surgical oncology in a state with a single-payer publically funded system.

Peri- and post-operative management

Patients were assessed at the same pre-operative clinic and managed on the same surgical ward by a common team of residents and fellows. All patients received antibiotic and deep venous thrombosis prophylaxis at induction of anaesthesia. Prophylactic octreotide was not used routinely. All patients received two (2) intra-abdominal closed-suction drains placed in proximity to the pancreatico-enteric and choledocho-enteric anastomosis. All patients received a nasogastric (NG) tube. Jejunal feeding tubes were only exceptionally used. PJ consisted of duct-to-mucosa, end-to-side anastomosis with a pancreatic duct stent and PG consisted of a 'dunking' single layer posterior-wall gastrostomy.[11] Routine testing of drain amylase was done in all cases but timing of drain removal was left to the treating team.

Outcomes definition

The ISGPF definition of a PF and both the ISGPF[12] and the Strasberg and Linehan[13] grading systems for PF severity were used. Delayed gastric emptying was assessed using the ISGPS definition.[14] Total 90-day complications were assessed using the Clavien–Dindo classification[15] as well as the Comprehensive Complication Index (CCI).[16] This tool is free to use and can be calculated online (http://www.assessurgery.com). Patients were considered ever-smoker if they were smoking at the time of surgery or had been smokers previously. Intra-abdominal collection was defined as the radiological or re-operative finding of fluid collections in the surgical field – no size or volume limitations were imposed. Wound infections were defined as any wound with cellulitis, purulent drainage, and any incision requiring opening and packing or vac therapy.

Statistical analysis

Statistical analysis was performed using the Stata software, version 12.1 (StataCorp, TX, USA; http://www.stata.com). *T*-tests and ANOVA were employed to compare the means of normally distributed continuous data, Wilcoxon's Rank-Sum test for non-normally distributed continuous data, and χ^2 test for categorical data. Propensity factors were identified *ad hoc* as variables associated with assignment to PG at univariate logistic regression with a $P < 0.15$. No propensity factors were pushed into the model *a priori*. The propensity score was obtained using a predictive function after a multivariate logistic regression with surgical technique as the outcome and the propensity factors as covariates. No quadratic functions or interaction terms were used to enhance the model. The propensity score was used as a co-variate in univariate and multivariate logistic regressions to determine the relationship between PG and surgical outcomes.

Results

Baseline characteristics and construction of the propensity score

One-hundred and three PG and 103 PJ were identified as having been performed by three surgeons. Table 1 summarizes baseline patient demographic, pre-operative and pathological characteristics. Both groups were relatively similar with regards to measured variables except for smoking, dyslipidemia, chronic obstructive pulmonary disease and diagnosis of intraductal papillary mucinous neoplasm (IPMN) or cholangiocarcinoma. These

Table 1 Patient characteristics in the pancreatico-gastrostomy (PG) and pancreatico-jejunostomy (PJ) groups

Variable	PG	PJ	*P*
Age (median and range)	66 (21–88)	66 (23–82)	0.95
Gender			
Men	55 (53.4)	46 (44.7)	
Women	48 (46.6)	57 (45.3)	0.78
BMI			
<25	44	43	
25–30	28	27	
30–35	7	7	
35–40	2	2	
>40	1	1	1.0
Ever-smoker	19	33	0.027
Hypertension	34	36	0.89
Dyslypidemia	20	12	0.096
Coronary artery disease	7	13	0.179
Diabetes	19	21	0.81
COPD	10	4	0.083
Peripheral vascular disease	3	4	0.73
ASA >3	48	38	0.106
Jaundice	49	51	0.76
Biliary drainage	40	38	0.8
Cholangitis	10	8	0.637
Weight loss >10%	19	15	0.45
Neoadjuvant chemotherapy	9	5	0.26
Neoadjuvant radiation	2	1	0.5
Tumour size (median, range)	3 (0.2–8.5)	3 (0.1–13)	0.66
T3	59	52	0.55
R1 resection	19	20	0.97
Portal vein resection	7	10	0.46
Lymph nodes (median, range)	10 (0–38)	9 (1–33)	0.9
Pathology			
Adenocarcinoma	46	43	0.7
Cholangiocarcinoma	10	2	0.018
Ampullary carcinoma	8	12	0.321
IPMN	11	3	0.029
Neuroendocrine	7	10	0.421
Duodenal	1	0	0.90
Non-malignant	8	9	0.80
Unknown	4	6	0.75
Others	8	18	

BMI, body mass index; COPD, chronic obstructive pulmonary disease; ASA, American Society of Anesthesiologists; IPMN, intraductal papillary mucinous neoplasm.

factors differed sufficiently between the two groups to be considered propensity factors and were used to construct the propensity score.

PF

Morbidity of the two groups is summarized in Table 2. Peri-operative transfusion and diagnosis of cholangiocarcinoma or adenocarcinoma had a statistically significant association with PF at univariate logistic regression and were included as risk factors in the multivariate analysis. A body mass index (BMI) of more than 35, although not statistically significant as a PF risk factor, was included in the multivariate analysis owing to a well-documented relationship between obesity, pancreatic texture and PF in the literature. Soft pancreatic texture [odds ratio (OR) = 21, P = 0.006] and pancreatic duct diameter [OR = 0.4, P = 0.027] were also statistically significant risk factors, but we could not be included in the multivariate analysis owing to missing data – including them as covariates radically decreased the sample size. Table 3a shows the results of univariate, propensity-score adjusted and multivariate propensity-score adjusted logistic regression examining the relationship between the reconstructive option and PF. The OR for PG was 1.19 (0.6–2.3) (P = 0.6) at univariate analysis, 0.87 (0.4–1.8) (P = 0.72) at propensity-score adjusted analysis and 0.97 (0.4–2.2) (P = 0.9) at propensity-score adjusted multivariate analysis.

DGE

At univariate, multivariate and propensity-score adjusted multivariate analysis, there was no effect of surgical reconstructive option on DGE grades B/C (Table 3.) There was a statistically significant difference in all-grade DGE (79/103 versus 55/103 for PG versus PJ respectively, P = 0.02). This remained significant at propensity-score adjusted multivariate regression (OR = 2.4, P = 0.012).

Other outcomes

The median length of stay was not statistically significant between the two groups (18 and 14 days in the PG and PJ groups, respectively, P = 0.20). There was no difference in all-cause 90-day mortality between the two groups (9/103 and 5/103, P = 0.27.) Six out of 14 deaths had a PH, 3 out of 14 had myocardial infarcts or cardiac arrest, 2 out of 14 complications of intra-abdominal sepsis not related to a PF, 2 out of 14 complications of angiographic interventions for bleeding and 1 out of 14 had multiorgan failure after anastomotic bleed causing multi-organ failure. The median CCI differed by 7 points between the two groups (P = 0.027) but this difference did not remain significant after adjusting for propensity score (P = 0.4.) The distribution of morbidity as measured by the Clavien–Dindo classification did not differ between the two groups (P = 0.85). CCI did not differ by propensity score (P = 0.4).

Table 2 Pancreatic fistula, delayed gastric emptying, morbidity and mortality in patients undergoing pancreatico-gastrostomy (PG) and pancreatico-jejunostomy (PJ) after a pancreaticoduodenectomy

Complication	PG	PJ	*P*
PF	23	20	0.45
ISGBP – grade A	0	1	
ISGPF – grade B	8	6	
ISGPF – grade C	15	13	0.67
PF – Strasberg & Linehan			
Grade I	0	1	
Grade II	8	4	
Grade IIIA	6	8	
Grade IIIB	2	1	
Grade IV	2	2	
Grade IVA	2	1	
Grade V	3	3	0.85
Death	9	5	0.27
Delayed gastric emptying	74/103	55/103	0.02
No DGE	29	48	
DGE Grade A	46	37	
DGE Grade B/C	28	18	0.6
ICU readmissions	12	16	0.415
TPN	41	40	0.89
Antibiotic therapy	76	58	0.011
CT/US procedure	36	28	0.23
Reoperation	8	13	0.25
Discharged with JP	12	9	0.45
Pulmonary complication	14	13	0.79
Thromboembolic complication	12	6	0.12
Intra-abdominal collection	28	26	0.64
Bleeding	17	12	0.28
Wound Infection	21	24	0.69
UTI	23	9	0.004
LOS (median)	18	14	0.20
Complications	69/103	62/103	
Clavien-Dindo I	3	8	
Clavien-Dindo II	28	20	
Clavien-Dindo IIIA	13	15	
Clavien-Dindo IIIIB	5	1	
Clavien-Dindo IV	5	5	
Clavien-Dindo IVA	6	8	
Clavien-Dindo V	9	5	0.285
CCI	38.4	31.4	0.02

PF, pancreatic fistula; DGE, delayed gastric emptying; ICU, intensive care unit; TPN, total parenteral nutrition; CT/US, computed tomography/ultrasound; UIT, urinary tract infection; LOS, length of stay.

Discussion

The proportion of patients developing PF in the current series (20.9%) is similar to that observed in other retrospective studies that use the ISGPF classification. We did not identify a significant effect of pancreatico-enteric reconstructive option on the odds of a PF even after adjusting for propensity score. There was more overall DGE with PG both at univariate and multivariate propensity-score adjusted analysis, but there was no increase in DGE of grades B/C and no impact on length of stay, Clavien–Dindo or CCI. The mortality rates in our series did not differ between the two groups. A significant proportion of deaths were related to PF, consistent with other published series.[17] The CCI and Clavien–Dindo grade of complications did not differ between the two groups.

The CCI is a recently published tool that measures the weighted cumulative impact of post-operative complications and adverse events on a continuous scale;[16] as such it is more sensitive than existing morbidity measures and better reflects the global morbidity experience of the patients. This is the first time the CCI has been used as a measure of morbidity aside from its validating publication.[16] The magnitude of the difference in CCI at univariate analysis is therefore difficult to interpret conclusively; the 7-point difference in CCI between the groups at univariate analysis is probably attributable to the increased delayed gastric emptying and slight increase in UTI requiring antibiotic treatment in the PG group. This finding highlights the sensitivity of the CCI as a measure of the total morbidity experience as compared with traditional measures such as the Clavien–Dindo.

The belief that PG is protective of PF was supported by a large number of observational trials; however, a meta-analysis by Wente *et al.* showed them to be clinically heterogeneous, lacking adjustment for known confounders and at high risk of bias;[7] moreover a strong publication bias in favour of PG was demonstrated. More recently published meta-analyses of randomized trials have not conclusively demonstrated PG's protective effect. The meta-analysis by Ma *et al.* of four randomized trials used a random-effects model and found more intra-abdominal collections with PJ (OR = 0.46, 95% CI 0.26–0.79, $P = 0.005$) but no difference in total morbidity, mortality, PF or DGE. Using the same four randomized trials and a less conservative fixed-effects model, a meta-analysis by Shen *et al.* found an increase in intra-abdominal complications, but no difference in total complications.[8] It remains unclear if particular subgroups, such as those with smaller pancreatic ducts, or a soft fatty pancreas, stand to gain from one technique compared with the other.

Our institution's setup provides a unique opportunity to study the outcomes between the two groups. This can contribute to the literature through the quasi-experimental process of patient allocation as well as the expertise-based nature of the interventions being compared; this mitigates the confoundedness that had been the main criticism of previously published retrospective and observational series showing superiority of PG. Moreover, referral

Table 3a Univariate, multivariate and propensity-score adjusted analysis of the effect of pancreatico-gastrostomy (PG) on pancreatic fistula (PF)

Outcome: PF	OR PG (95% CI)	OR (95% CI) Cholangiocarcinoma	OR (95% CI) Adenocarcinoma	OR (95% CI) BMI>35	OR (95% CI) Perioperative-Transfusion	OR (95% CI) Propensity score
Univariate	1.19 (0.6–2.3) (*P* = 0.6)	5.9 (1.7–19) (*P* = 0.004)	0.4 (0.2–0.9) (*P* = 0.026)	1.7 (0.3–9.8) (*P* = 0.53)	2.1 (1.03–4.64) (*P* = 0.041)	3.8 (0.4–29) (*P* = 0.2)
Propensity-score adjusted	0.87 (0.4–1.8) (*P* = 0.72)					4.4 (0.5–40) (*P* = 0.18)
Multivariate and propensity-score adjusted	0.97 (0.4–2.2) (*P* = 0.9)	14.3 (2.2–94) (*P* = 0.006)	0.44 (0.18–1.1) (*P* = 0.069)	2.3 (0.4–15) (*P* = 0.4)	2.2 (0.9–5.4) (*P* = 0.14)	0.3 (0.02–5) (*P* = 0.4)

Table 3b Univariate, multivariate and propensity-score adjusted analysis of the effect of pancreatico-gastrostomy (PG) on delayed gastric emptying (DGE)

Outcome: DGE B/C	OR PG (95% CI)	OR (95% CI) Adenocarcinoma	OR (95% CI) Neuroendocrine	OR (95% CI) Jaundice	OR (95% CI) Propensity Score
Univariate	1.3 (0.6–2.6) (*P* = 0.6)	0.3 (0.13–0.7) (*P* = 0.005)	3.3 (1–10) (*P* = 0.05)	0.4 (0.2–0.95) (*P* = 0.04)	14.3 (1.5–133) (*P* = 0.02)
Propensity-score adjusted	0.94 (0.4–2.2) (*P* = 0.9)				14.9 (1.4–160) (*P* = 0.025)
Multivariate and propensity-score adjusted	1.02 (0.4–2.2) (*P* = 0.9)	0.4 (0.15–1.1) (*P* = 0.078)	1.8 (0.4–7) (*P* = 0.3)	0.6 (0.2–1.5) (*P* = 0.27)	10.8 (0.8–136) (*P* = 0.064)

bias is at least partially adjusted for by the propensity score. This might explain why our results are more consistent with the conclusions drawn from most of the randomized studies rather than the observational series comparing PG and PJ.

We chose to adjust for the propensity score instead of matching upon it. Similar conclusions are reached when we matched upon propensity score; however, this approach resulted in a decrease in our sample size owing to unmatched individuals. Using the propensity score as a co-variate in multivariate adjustment is as valid as its use to match individuals;[18,19] moreover, matching in an observational dataset is not without its caveat as it can be a source of bias itself.

In addition to the quasi-experimental setting, the presence of a strong surgeon preference for PJ or PG makes the study akin to an expertise-based trial, with each technique performed well beyond the learning curve of surgeons. A notable strength of our study's design lies in its comparison of two techniques performed by expert surgeons well beyond their procedural learning curves. Effect measures from expertise-based comparisons can therefore more accurately reflect effects seen in clinical settings than trials where procedures are compared at different levels of proficiency.[10]

The possibility of residual confounding remains a concern given our inability to adjust for pancreatic duct diameter and pancreatic texture; however, our adjustment for BMI and pathology could have at least partially adjusted for these confounders. We were unable to perform subgroup analysis to identify whether PG had more benefit in a particular patient subpopulation, but we can postulate that outcomes are similar in the subset of patients with a soft pancreatic texture or a small duct.

With no dramatic differences in surgical outcomes at our site, this study would suggest that pancreatic surgeons should continue using the reconstructive technique they are most familiar with and ones that have yielded institutional outcomes similar to those in published series.

Acknowledgements

The authors would like to thank Dr Slankamenac for kindly providing us with the Comprehensive Complication Index score for our patients, and Marcel Edwards and Yasmine Youssef-Elhaili for their help with data collection.

Conflicts of interest

None declared.

References

1. Siegel R, Naishadham D, Jemal A. (2012) Cancer statistics, 2012. *CA Cancer J Clin* 62:10–29.
2. Ouaissi M, Giger U, Louis G, Sielezneff I, Farges O, Sastre B. (2012) Ductal adenocarcinoma of the pancreatic head: a focus on current diagnostic and surgical concepts. *World J Gastroenterol* 18:3058–3069.
3. Poon RT, Lo SH, Fong D, Fan ST, Wong J. (2002) Prevention of pancreatic anastomotic leakage after pancreaticoduodenectomy. *Am J Surg* 183:42–52.
4. Kleespies A, Rentsch M, Seeliger H, Albertsmeier M, Jauch KW, Bruns CJ. (2009) Blumgart anastomosis for pancreaticojejunostomy minimizes severe complications after pancreatic head resection. *Br J Surg* 96:741–750.
5. Peng SY, Wang JW, Lau WY, Cai XJ, Mou YP, Liu YB *et al*. (2007) Conventional versus binding pancreaticojejunostomy after pancreaticoduodenectomy: a prospective randomized trial. *Ann Surg* 245:692–698.

6. McKay A, Mackenzie S, Sutherland FR, Bathe OF, Doig C, Dort J *et al*. (2006) Meta-analysis of pancreaticojejunostomy versus pancreaticogastrostomy reconstruction after pancreaticoduodenectomy. *Br J Surg* 93:929–936.
7. Wente MN, Shrikhande SV, Muller MW, Diener MK, Seiler CM, Friess H *et al*. (2007) Pancreaticojejunostomy versus pancreaticogastrostomy: systematic review and meta-analysis. *Am J Surg* 193:171–183.
8. Shen Y, Jin W. (2012) Reconstruction by Pancreaticogastrostomy versus Pancreaticojejunostomy following Pancreaticoduodenectomy: a Meta-Analysis of Randomized Controlled Trials. *Gastroenterol Res Pract* 2012:627095.
9. Wellner U, Brett S, Bruckner T, Limprecht R, Rossion I, Seiler C *et al*. (2012) Pancreatogastrostomy versus pancreatojejunostomy for RECOnstruction after partial PANCreatoduodenectomy (RECOPANC): study protocol of a randomized controlled trial UTN U1111-1117-9588. *Trials* 13:45.
10. Devereaux PJ, Bhandari M, Clarke M, Montori VM, Cook DJ, Yusuf S *et al*. (2005) Need for expertise based randomised controlled trials. *BMJ* 330:88.
11. Zenilman MEMDFACS. (2000) Use of pancreaticogastrostomy for pancreatic reconstruction after pancreaticoduodenectomy. *J Clin Gastroenterol* 31:11–18.
12. Bassi C. (2005) Postoperative pancreatic fistula: an international study group (ISGPF) definition. *Surgery* 138:8–13.
13. Strasberg SM, Linehan DC, Clavien PA, Barkun JS. (2007) Proposal for definition and severity grading of pancreatic anastomosis failure and pancreatic occlusion failure. *Surgery* 141:420–426.
14. Wente MN, Bassi C, Dervenis C, Fingerhut A, Gouma DJ, Izbicki JR *et al*. (2007) Delayed gastric emptying (DGE) after pancreatic surgery: a suggested definition by the International Study Group of Pancreatic Surgery (ISGPS). *Surgery* 142:761–768.
15. Clavien PA, Barkun J, de Oliveira ML, Vauthey JN, Dindo D, Schulick RD *et al*. (2009) The Clavien-Dindo classification of surgical complications: five-year experience. *Ann Surg* 250:187–196.
16. Slankamenac K, Graf R, Barkun J, Puhan MA, Clavien PA. (2013) The comprehensive complication index: a novel continuous scale to measure surgical morbidity. *Ann Surg* 258:1–7.
17. Addeo P, Delpero JR, Paye F, Oussoultzoglou E, Fuchshuber PR, Sauvanet A *et al*. (2013) Pancreatic fistula after a pancreaticoduodenectomy for ductal adenocarcinoma and its association with morbidity: a multicentre study of the French Surgical Association. *HPB* 6:46–55.
18. Rosenbaum PR, Rubin DB. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41–55.
19. Senn S, Graf E, Caputo A. (2007) Stratification for the propensity score compared with linear regression techniques to assess the effect of treatment or exposure. *Stat Med* 26:5529–5544.